Chem Soc Rev

This article was published as part of the

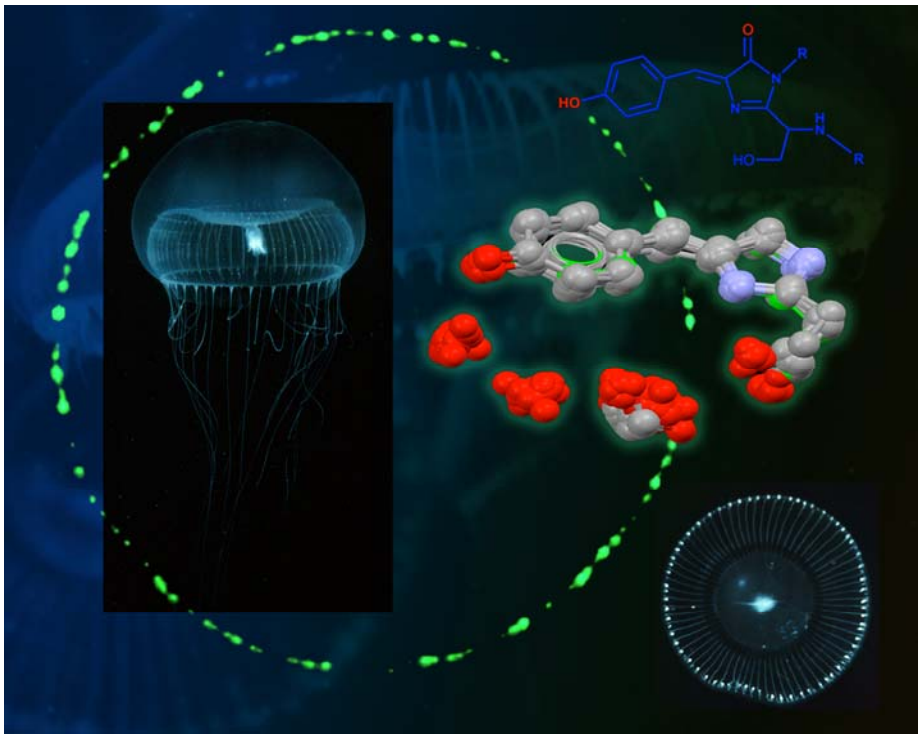
2009 Green Fluorescent Protein issue

Reviewing the latest developments in the science of green fluorescent protein

Guest Editors Dr Sophie Jackson and Professor Jeremy Sanders

All authors contributed to this issue in honour of the 2008 Nobel Prize winners in Chemistry, Professors Osamu Shimomura, Martin Chalfie and Roger Y. Tsien

Please take a look at the issue 10 table of contents to access the other reviews



The folding, stability and conformational dynamics of β-barrel fluorescent proteins†

Shang-Te Danny Hsu, Georg Blaser and Sophie E. Jackson*

Received 28th July 2009
First published as an Advance Article on the web 4th September 2009
DOI: 10.1039/b908170b

This *critical review* describes our current knowledge on the folding, stability and conformational dynamics of fluorescent proteins (FPs). The biophysical studies that have led to the elucidation of many of the key features of the complex energy landscape for folding for GFP and its variants are discussed. These illustrate some important issues surrounding how the large β-barrel structure forms, and will be of interest to the protein folding community. In addition, the review highlights the importance of some of these results for the use of FPs in *in vivo* applications. The results should facilitate and aid in experimental designs of *in vivo* applications, as well as the interpretation of *in vivo* experimental data. The review is therefore of interest to all those working with FPs *in vivo* (103 references).

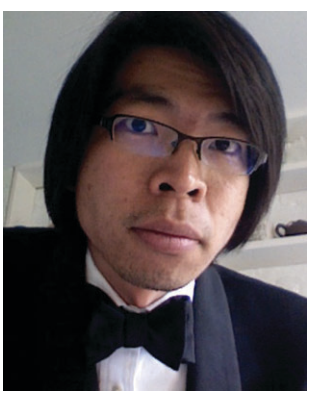
1. Introduction

It has long been established that GFP and other FPs first have to fold to a near-native like structure before the chromophore can form by a cyclisation and oxidation of the polypeptide backbone[1] (for details of the role of the folding and structure of GFP in chromophore maturation see the review by Craggs in this special issue[2]). In addition to this, it has been shown that the protein structure itself plays an essential role in creating and maintaining a semi-rigid environment around the chromophore where bulk solvent molecules are excluded and the conformational flexibility of the chromophore is low (Fig. 1). This greatly reduces the fluorescence deactivation pathways and leads to a state with a high quantum yield (for further details on this, see the review by van Thor in this special issue[3]). In this Review, we focus on biophysical studies on FP stability, folding and conformational dynamics and illustrate how these relate to the function of FPs in biological applications.

The Review starts with a description of how protein engineering techniques have been used to improve the folding properties of many FPs using random mutagenesis and selection methods (section 2). A wide range of different techniques have been employed to study the stability of the native states of FPs *in vitro*, including denaturation by acid, pressure, mechanical force and chemical denaturants. The results of these studies and what has been learnt about FP stability are described in detail in section 3. Not only have

Chemistry Department, University of Cambridge, Lensfield Road, Cambridge, UK CB2 1EW. E-mail: sej13@cam.ac.uk; Fax: +44 1223 336362; Tel: +44 1223 760211

† Part of a themed issue on the topic of green fluorescent protein (GFP) in honour of the 2008 Nobel Prize winners in Chemistry, Professors Osamu Shimomura, Martin Chalfie and Roger Y. Tsien.



Shang-Te Danny Hsu

Dr Shang-Te Danny Hsu is a Research Fellow at Wolfson College in Cambridge University. He received his BA and MSc from the National Tsing Hua University in Taiwan. In 2004, he completed his PhD with honour in Utrecht University, the Netherlands. He was awarded a Netherlands Ramsay Fellowship from the Royal Netherlands Academy of Arts and Sciences and subsequently a Human Frontier Science Program Long-term Fellowship to study co-translational protein folding by NMR spectroscopy with Prof. Chris Dobson FRS in Cambridge, which led to a Young Investigator Award from the International Society of Magnetic Resonance in 2007.

Georg Blaser

Dr Georg Blaser is a Postdoctoral Research Associate in the Chemistry Department at Cambridge University. He received his doctorate in Bioorganic Chemistry from Durham University and his Master's Degree in Chemistry from Bristol University, UK. His current research focuses on folding dynamics and thermodynamical stability measurements of proteins using single-molecule FRET and ensemble measurements.

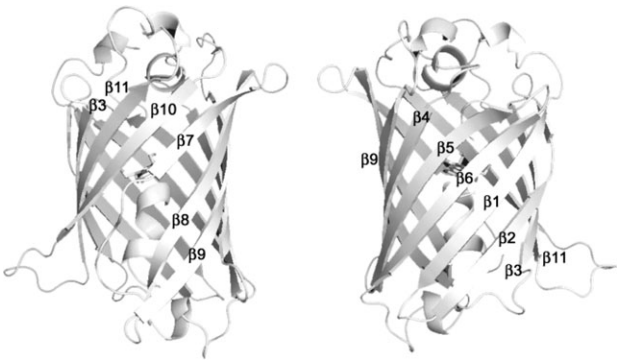


Fig. 1 Schematic representations for the structure of GFPuv (PDB code: 1B9C) drawn by PyMol (DeLano Scientific LLC), viewed from two opposite sides. The chromophore is shown in stick mode. Each β-strand is numbered from the N to the C terminus.

these studies shed light on some of the factors that govern the stability of this class of proteins, but they have revealed complex energy landscapes with unique features which are also of importance to the protein folding field. Section 4 reports on the studies which have been undertaken over the past decade to understand the unfolding and folding pathways of FPs and GFP in particular. Again, a complex energy landscape has been uncovered and some of the rate-limiting steps in the folding of the complex β-barrel structure determined. The single-molecule methods that have been employed and which have generated further details on the unfolding pathways of GFP are summarised in section 5. Despite the complexity of the GFP structure, a great deal is now known about the critical steps in folding of this protein.

Section 6 of the Review draws upon the knowledge gained from *in vitro* measurements of folding and stability, and discusses the implications for some applications of FPs *in vivo*. The ability of fragments of GFP to form stable elements of structure which are not rapidly degraded *in vivo* and the application of these to split-GFP technology are discussed in section 7. Two powerful techniques arise as a direct result of being able to divide the GFP structure into complementary fragments which can reassociate *in vivo* to generate a fluorescent state: the split-GFP solubility reporter assay and the so-called biomolecular fluorescence complementation (BiFC) method. Finally, the use of GFP in probing the mechanisms by which proteins are unfolded and degraded by cellular degradation machines such as the Clp system of prokaryotes and the proteasome in eukaryotes is described.

Sophie E. Jackson

Dr Sophie Jackson is a Reader in Biophysical Chemistry at Cambridge University. She completed a B.A. in Chemistry at Oxford University before starting a PhD at Imperial College, University of London which was completed in Cambridge. She was awarded a William Stone Research Fellowship (Peterhouse, Cambridge) and a Human Frontiers Science Program Organisation Postdoctoral Fellowship in the Chemistry Department at Harvard University. She returned to Cambridge in 1995 as a Royal Society University Research Fellow. Her research focuses on different aspects of protein folding and assembly processes, and her current interests include understanding how large complex protein structures form.

2. Improving the folding efficiency of GFP: protein engineering studies

It has been known for some time that wild-type GFP is prone to misfolding and aggregation leading to reduced chromophore maturation and low yields.[4] In addition to this intrinsic tendency of GFP to misfold and aggregate, the situation is worse when GFP is fused to other relatively insoluble proteins,[5] or when using circular permutants of GFP which are often employed as biosensors.[6–8] It's not surprising, therefore, that a large number of studies have focused on improving the folding properties of GFP and other FPs, and almost all of these have taken advantage of the intrinsic fluorescence of the protein to act as a marker of efficient folding, chromophore maturation and solubility. Although a number of different approaches have been taken, they all in some way use random mutagenesis in conjunction with selection methods based on cellular fluorescence to identify mutations in GFP and other FPs which produce brighter proteins *in vivo*. Many of the mutants discovered using this strategy are termed "folding mutants", however, it is clear that some of them act by improving the efficiency of chromophore formation and not folding *per se*. These mutations are not considered here but are discussed in detail in the review by Craggs on GFP chromophore maturation.[2] The results from a number of different laboratories, all of which have succeeded in improving the folding properties of GFP and other FPs, are discussed below.

2.1 Green fluorescent protein

Some of the earliest studies aimed at improving the folding properties of fluorescent proteins were on GFP. Using random mutagenesis and screening for increased brightness of colonies in *E. coli*, the Haseloff group identified some of the first "folding mutants", Val→Ala163 and Ser→Gly175, which gave rise to a 35-fold increase in green fluorescence intensity in both *E. coli* and yeast at 37 °C.[9] One of these mutations, Val→Ala163, was also found independently in later studies. Using a similar approach, the Kohno group also identified another mutation, Ser→Pro147, which showed increased fluorescence at elevated temperatures.[10]

One of the most important early studies was conducted by the Stemmer group who used an alternative approach, DNA shuffling, to generate a library of mutants of GFP from which to select. In their case, a mutant 42 times more fluorescent than wild type was identified which contained three mutations, Phe→Ser99, Met→Thr153 and Val→Ala163.[11] It was named the "cycle 3" mutant but is now more commonly referred to as

GFPuv. For this mutant, an *in vitro* study of the folding kinetics was undertaken in order to understand the origin of the improved brightness and fluorescence observed *in vivo*. Initial studies, somewhat surprisingly, revealed that the folding kinetics of the cycle 3 mutant are very similar to wild type,[12] however, a more detailed study established that the mutations greatly reduced aggregation and thereby resulted in more efficient folding and higher yields.[13] It is thought that the mutations, which all lie on the same face of the β-barrel structure, act by reducing the overall hydrophobicity of GFP, thereby suppressing aggregation.[12,13] The three mutations in the cycle 3 variant of GFP have also been used in conjunction with other mutations known to improve chromophore maturation (Phe → Leu64 and Ser → Thr65) to produce a GFP, known as *gfp+*, which exhibited up to a 320-fold increase in detection level and which has been shown to be an accurate reporter of gene expression levels over several orders of magnitude when used in *in vivo* gene expression assays.[14]

In a more recent study, DNA shuffling techniques were also employed to generate a library of GFP mutants. In this case, the cycle 3 mutant of GFP was fused to a very poorly folding bait protein, bullfrog red cell H subunit of ferritin, which is insoluble when expressed at 37 °C.[15] This so called "folding interference method" was used and a "superfolder" GFP identified after four rounds of selection, which contained six additional mutations to the three already incorporated into the parent cycle 3 protein. For the superfolder GFP, a comprehensive investigation of the effects of each of the six mutations was undertaken in order to understand how each promoted folding and chromophore maturation. Two of the mutations, Ser → Arg30 and Tyr → Asn39, were found to increase the stability and folding rate of GFP, and a crystal structure showed that there is a reorganisation of the side chains around the sites of mutation which results in a more favourable network of electrostatic interactions and hydrogen bonding. In contrast, two of the mutations (Tyr → Phe145 and Ile → Val171) had little effect on the folding rate or stability of GFP, and it has been assumed that they act by reducing aggregation. The other two (Asn → Thr105 and Ala → Val206) also have no effect on folding or stability but it was noted that they increase the β-propensity of the polypeptide chain.

2.2 Other fluorescent proteins (FPs)

Mutations which improve the folding and fluorescence of GFP have also been incorporated into CFP and YFP to improve the folding properties of these two proteins.[16] Monomeric, optimised versions of eCFP and eYFP, which fold faster and more efficiently at 37 °C and which had superior solubility and brightness, have been produced. Improvements were most pronounced in *E. coli* but improvements in mammalian cell lines were also observed. In another study, the Gadella group used a similar approach but, in this case, used mutations which had been shown to improve folding/fluorescence of eYFP and eCFP, and introduced them into eGFP and eBFP to create SGFP2 and SBFP2, strongly enhanced versions of GFP and BFP.[17]

3. Stability measurements

Over the past decade, a range of different methods have been used to study the stability of the native state of GFP, its variants and FPs in general. In this section, stability refers to thermodynamic stability, that is, the difference in the Gibbs free energy between the native and denatured states. Fluorescent proteins have been denatured with acid,[18–20] pressure,[21] temperature,[21] force[22–26] and chemical denaturants[13,27–33] and both ensemble and single-molecule measurements have been made. The unfolding of FPs has been followed using different optical probes including the fluorescence of the chromophore and tryptophan/tyrosine residues, circular dichroism, UV/Vis absorbance, FT-IR, atomic force microscopy (AFM), in addition to nuclear magnetic resonance (NMR) spectroscopy including ^{15}N, ^{19}F and photoCIDNP experiments.[19,34-37] The next sections describe in detail what has been learnt about the stability and unfolding of FPs from these studies.

3.1 Chemical denaturation

One of the most extensively used methods for studying the stability and unfolding of proteins under equilibrium conditions is chemical denaturation using reagents such as urea, guanidinium chloride (GdmCl) and guanidinium thiocyanate (GdmSCN).[38] These methods have been been applied to FPs from both *Aequorea* and *Anthozoa*.

Early studies by the Kuwajima group on wild-type GFP and the cycle 3 mutant used GdmCl to unfold the proteins and established that they were both very stable against chemical denaturation, with the cycle 3 mutant apparently being more stable than wild type.[13] However, their results also showed that the unfolding equilibrium was not fully reached even after several days,[13] unusual behaviour not observed for the unfolding of most proteins. The Uversky and Turoverov groups have studied the stability of FPs with different oligomeric states (including monomeric EGFP and mRFP1, dimeric "dimer 2", and tetrameric zFP506 and DsRed) using GdmCl, and also observed that the systems were slow to reach equilibrium.[27] In this case apparent thermodynamic stabilities of the FPs studied were calculated and ranged from 4.6 to 10.6 kcal mol^{-1}. Although the tetrameric zFP506 was found to be the most stable, oligomeric state was not the only factor that was found to influence the stability (Fig. 2A).[27]

These chemical denaturation studies have now been followed up by work from several other groups who have probed in further detail the GdmCl-induced unfolding of GFP variants under pseudo-equilibrium conditions. Both our own group and the Jennings group have shown a time-dependent shift in the position of the unfolding equilibrium for a truncated form of GFPuv (trGFPuv)[31] and the superfolder GFP (sfGFP)[30] (Fig. 2B and C). We found that trGFPuv did not reach equilibrium for over two months[31] (Fig. 2B). A careful analysis of the results from these experiments revealed that an intermediate state was populated under the conditions used, and that this state was both stable (with a free energy of unfolding of 6.0 kcal mol^{-1} with respect to the denatured state) and had considerable structure (with an *m* value of 4.2 kcal mol^{-1} M^{-1} with respect to the denatured state;

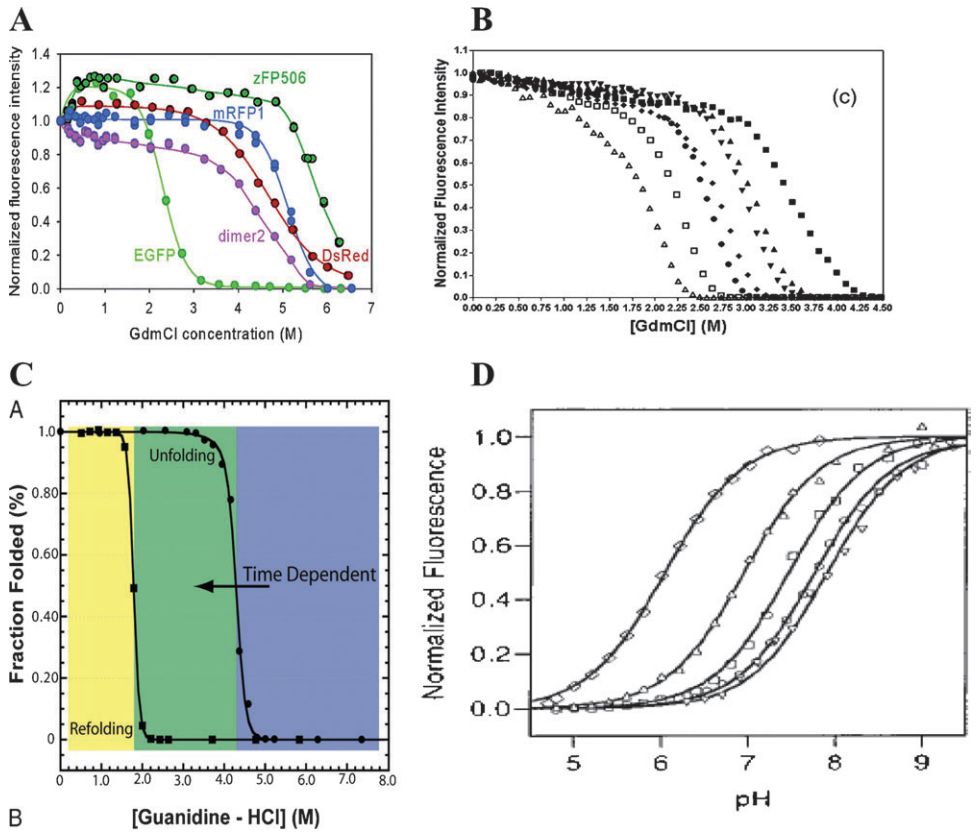


Fig. 2 Pseudo-equilibrium curves for the unfolding of GFP variants under varying conditions. (A) Unfolding curves of EGFP, zFP506, mRFP1, dimer 2, and DsRed after 9 days (adapted from ref. 27). (B) Green fluorescence of trGFPuv measured as a function of GdmCl concentration at pH 6.0 at 25 °C as a function of equilibration time: from right to left 3 h, 12 h, 24 h, 48 h, 5 days, 13 days, 44 days (adapted from ref. 31). (C) The non-coincident equilibrium unfolding (●) and refolding (■) transitions of sfGFP induced by GdmCl at pH 6.8 at 25 °C after 96 hours showing the hysteresis. The unfolding transition exhibits time dependence and collapses toward the refolding transition over months. Lines are fitted to a two-state model for refolding and a three-state model for unfolding and different zones of transition curves are coloured yellow, green and blue (adapted from ref. 46). (D) pH dependence of the chromophore fluorescence of various YFP variants (from left to right the mutants were: S65T, wt, H148Q, S65T/H148Q, H148G). Data points were fitted to a theoretical titration curve with one titratable group.

m values are a measure of the change in solvent accessible surface area between two states[38]). In fact, the stability of the intermediate state was comparable to that observed for many small monomeric proteins.[39] In addition, far UV-CD and tryptophan/tyrosine fluorescence measurements established that the intermediate state contained considerable secondary and tertiary structure whilst its green fluorescence was quenched to about 10% of that found in the native state.[31] The quenching of the green fluorescence in the intermediate state was attributed to a local unfolding of β-strands 7–9 resulting in the partial opening of the β-barrel structure and exposure of the chromophore to solvent. This was supported by evidence from NMR H/D exchange experiments from both our own[31] and other groups,[40] see section 3.4 for further details.

The Jennings group have also studied the GdmCl-induced unfolding of another GFP variant, superfolder GFP (sfGFP). They too observed time-dependent shifts in the unfolding curves, and also undertook refolding experiments. These established that there was hysteresis between the equilibrium curves starting from denatured or native states of the protein (Fig. 2B).[30] In this case, they performed a comprehensive analysis using mutants that were incapable of forming the chromophore and determined that the hysteresis was due to the chromophore itself which creates a rough energy landscape on which the protein unfolds/folds.[30] Further details are given in section 3.5.

More recently, other groups have also published denaturant studies and confirmed the population of an unfolding intermediate under equilibrium conditions. Wielgus-Kutrowska, Clark and colleagues demonstrated that a partially unfolded intermediate with molten-globule like properties is present in the GdmCl-induced unfolding of S65T/G67A GFP, a mutant which is unable to form the chromophore and which showed reversible unfolding in chemical denaturants.[28] Xie and Zhou have also observed three-state behaviour for the GdmCl-induced unfolding of GFPuv and shown that folding and unfolding proceed in a stepwise manner *via* a compact but fluorescent intermediate state that is populated under equilibrium conditions.[29]

Further evidence for an intermediate state has come from studies on the unfolding of a variant of YFP (Citrine), and has been demonstrated using both single-molecule and ensemble experiments.[33] Here, the fluorescence of the intrinsic green/yellow chromophore along with that from a covalently attached dye (Alexa647) was used to determine the unfolding

behaviour of the protein at a single molecule level. Single-pair fluorescence resonance energy transfer (sp-FRET) and two-colour coincidence detection (TCCD) techniques were employed. Results from both the single-molecule experiments and ensemble measurements showed biphasic equilibrium denaturation curves indicating the presence of a stable intermediate state for this modified form of Citrine.[33]

3.2 Acid denaturation

Many proteins can be reversibly denatured by acid, a result of the effect of perturbed pK_a values for the side chains of Glu and Asp residues in the native states of many proteins, together with unfavourable electrostatic interactions due to the build up of positive charge on most proteins at low pH.[38] The acid-unfolding of several different FPs has been studied, with an early study by Vrzheshch and coworkers on the denaturation of tetrameric DsRed under mildly acidic conditions.[20] A loss of fluorescence was observed on lowering the pH to 4–4.8 which was shown to be associated with a partial unfolding of the structure. Re-alkalization of the solution resulted in only a partial recovery of structure and fluorescence, indicating that, for DsRed, the acid unfolding was not fully reversible.[20]

The Kuwajima group has used acid-denaturation together with a wide range of spectroscopic and other probes, to characterise in detail the pH-unfolding of GFPuv. Similar to DsRed, they observed acid-unfolding events at pH values below 5.[18] Although the focus of this work was on the refolding kinetics of GFPuv from the acid-unfolded state (see section 4 for further details), it was followed up with a detailed study of the denatured state of GFPuv at pH 4.0.[41] Small-angle X-ray scattering and fluorescence measurements revealed that the protein adopts a partially structured intermediate state under these conditions. This state was found to be compact with respect to the denatured state, however, still expanded relative to the native state (by some 40%). Although the hydrophobic core around the single tryptophan residue at position 57 was found to be largely formed, the tertiary structure around the green chromophore was not fixed, hence the lack of green fluorescence of this state. The intermediate state showed some of the characteristics displayed by molten-globule states. Kinetic refolding experiments on this species established that it had properties similar to one of the main intermediates transiently populated during kinetic refolding experiments.[18]

Other techniques have also been applied to obtain information on the acid-denatured state of GFPuv. ^{19}F-NMR spectroscopy in combination with photochemically induced dynamic nuclear polarisation (CIDNP) techniques has been used to study both the acid- and GdmCl-induced unfolding of trGFPuv whose tyrosine residues had been selectively ^{19}F-labelled.[19] Although there was no detectable difference in the ^{19}F-spectra of the two different denatured states, photo-CIDNP experiments and far-UV CD spectra clearly showed evidence for significant residual structure at pH 2.9. This was in contrast to the results obtained in high concentrations of chemical denaturant (6 M GdmCl) or at lower pH (1.5), where there was little evidence of structure. In particular, the pH 2.9 state was shown to be a mixture of highly flexible and more compact structures as indicated by an analysis of the nuclear and electron spin relaxation processes.[34] More recently, a pulsed-labelling H/D exchange NMR strategy was used to probe in detail the residual structure at pH 2.9.[32] These experiments provided evidence that the structure was non-native like and it was shown to reside almost exclusively in β-strands 1 and 3. It has been suggested that this structure is stabilised by a local hydrophobic collapse.[32]

Recently, our laboratory has been studying the pH stability of the YFP Venus and found that it can be influenced dramatically by pH. On lowering the pH from 8.0 to 6.0, the midpoint for unfolding decreased significantly from 5.8 to 3.6 M GdmCl.[42] We also have evidence that small scale changes in structure that occur in Venus at lower pH values (pH 6.0) result in increased sensitivity of the yellow chromophore towards environmental conditions such as ionic strength and chloride ions.[42]

3.3 Pressure denaturation

Although not as commonly used as chemical denaturants or acid, pressure can also be employed to unfold proteins.[43] Such an approach has been used on a red-shifted variant of GFP (rsGFP) which was found to be very resistant to pressure denaturation at room temperature, withstanding pressures of up to 9 kbar.[21] However, at elevated temperatures, rsGFP was observed to undergo both partial and full pressure denaturation. At 58 °C, a co-operative collapse of the β-can structure (as probed by FT-IR measurements of secondary structure) corresponding to a large-scale unfolding event was observed at 7.8 kbar. A large change in volume of about 250 ml mol^{-1} was associated with this transition. Interestingly, the pressure-induced denatured state was found to be different to that observed in thermal denaturation experiments, and α-helical elements of structure in particular seem to be resistant to pressure denaturation. At around 4.5 kbar, there was some evidence for another, much smaller, structural change and this was investigated further under different conditions. At 5 and 30 °C, this transition could be seen more clearly and shown to be associated with a small conformational change in the region of the chromophore with no major changes in secondary structure. It was proposed that this transition may involve the penetration of water molecules into the β-can structure resulting in the quenching of the chromophore fluorescence. The state formed at 4 kbar at 30 °C has been termed a 'swollen pre-transitional' state and had a relatively small free energy change of $\sim$10 kJ mol^{-1} associated with it.

Although high temperatures are commonly used to study protein denaturation processes, the thermal unfolding of many FPs including rsGFP is not reversible with the protein aggregating at high temperatures.[21] As a result of this, this technique has not been widely employed to study the stability of FPs.

3.4 NMR experiments: H/D exchange and conformational dynamics

The H/D exchange of amide protons in conjunction with multi-dimensional NMR spectroscopy has been used

extensively to study non-native states in proteins.[44] Such studies provide information on partially structured, high energy states on the energy landscape for folding. These approaches can also be used to investigate residual structure in denatured states. These techniques have been used by several groups to study the denatured and partially structured states of FPs, particularly GFPuv.

The H/D exchange rates of amide protons in ^{15}N-labelled samples of GFPuv or trGFPuv have been reported in two separate studies. The earlier study by the Holak group established that, although the β-barrel structure of GFP was rigid on the picosecond to nanosecond time scale as probed by ^{15}N-relaxation techniques, there was conformational flexibility and motions on the micro- to millisecond time frame in β-strands 3, 7, 8 and 10, as shown by increased H/D exchange rates.[40] In addition, they showed that His148 affects the conformational stability of GFPuv and the spectra of a mutant (H148G) showed double backbone amide resonances typical of two conformations in slow exchange. Similar results were also found in H/D exchange studies on trGFPuv, with higher rates observed for amides in β-strands 7, 8, 9 and 10 indicating conformational flexibility in this region[31] (Fig. 3a and b). In this study, exchange rates were measured over extended periods of time and 40 amide protons were found to be protected from exchange even after several months (Fig. 3a). These groups clustered to form a stable core region encompassing most of the β-strands and located at one end of the barrel structure[31] (Fig. 3c), and it was proposed that these residues play an important role in stabilising the intermediate state of GFPuv. Further, it was shown that the exchange rates of these amide groups were not dependent upon the concentration of denaturant (between 0–1 M GdmCl), suggesting that exchange in this region does not occur through a local unfolding mechanism but instead is dominated by solvent penetration.[32] This work was followed up with an additional study which used higher pH, temperature and GdmCl to accelerate the rate of amide exchange, thereby enabling even the slowest of exchanging protons to be studied.[32] Even in this case, some of the amide protons did not undergo exchange even after several months of incubation.[32] These included the amide protons of residues Ile14, Leu15, Tyr92, Tyr106, Val112, Lys113, and Asn120 which cluster together to form a super-stable core located in β-strands 1, 4, 5 and 6 (Fig. 4A).

H/D exchange was also used to identify residual structure in the acid-denatured state of GFPuv (pH 2.9).[32] A pulsed-labelling strategy was employed and exchange in the low pH state monitored. Even after long incubation periods (80 min), amide protons of residues Val12, Ile14, Leu15, Val16, Glu17, Leu18, Leu44, Lys45, Phe46, and Ile47 were found not to have undergone exchange, indicating that they were protected and suggesting that they are involved in residual structure located in β-strands 1 and 3 (Fig. 4B). These residues are mainly hydrophobic in nature leading to the suggestion that the residual structure may be stabilised by a local hydrophobic collapse occurring in the N-terminal region of the protein.[32]

More recently, H/D exchange studies on YFP Venus have shown the most protected set of amide residues observed to date.[42] Even at pH 8.0 and 37 °C, some amide protons in Venus remained clearly visible in the corresponding ^{15}N–^{1}H correlation spectrum even after more than 13 months. This extraordinary stability, seen for a number of residues in β-strands 4–6, 10 and 11, provides further evidence that the β-barrel structures of FPs contain a super-stable core.[42]

In addition to the H/D exchange studies, NMR relaxation techniques have been used to obtain information on the conformational dynamics of FPs. In 2002, Seifert and coworkers published an NMR study on EGFP and ECFP in which fluorinated tryptophans were incorporated into the proteins and ^{19}F-NMR techniques were employed to investigate slow molecular motion in the proteins.[45] A slow exchange process between two states was observed on the millisecond time scale and attributed to conformational hetereogeneity of either the green chromophore or His148.[45]

3.5 Hysteresis in GFP unfolding and refolding

Hysteresis literally translates as "lagging behind". It describes a property of systems which fail to respond to an applied external force immediately but rather remain in their original state for a so-called lag phase before reaching a final state. Hysteresis has been observed in the folding and unfolding of sfGFP resulting in titrations with GdmCl starting from unfolded and folded sfGFP which were non-coincident[30] (Fig. 2C). Under highly denaturing or very native-like conditions, a typical transition from native to denatured state was observed (Fig. 2D). However, the unfolding transition was time-dependent and collapsed toward the refolding transition over months as had been observed in other studies (see section 3.1). At high denaturant concentrations ([GdmCl] $>$ 5 M), unfolding appears to progress predominantly from the native state ensemble (N_{nat}) to the unfolded state (U). At low denaturant concentrations ([GdmCl] $<$ 1.5 M), refolding seems to follow two-state behaviour from U directly to N_{nat}. However, at intermediate denaturant concentrations, a native-like intermediate state termed N_{iso} is proposed to form a kinetic trap resulting in the hysteresis. Consequently, exchange between the native state N_{nat} and native-like intermediate (N_{iso}) is slow. Experiments with a mutant incapable of forming the chromophore demonstrated that the chromophore itself is responsible for the hysteresis. Once the chromophore has formed, the energy landscape for folding for sfGFP becomes rough, in contrast to the smoother energy landscape for the *de novo* folding of GFP which has yet to undergo the post-translational modification and chromophore formation.[30]

The nature of N_{iso} and origins of the hysteresis were investigated further by combining experimental (NMR) and computational methods.[46] A coarse-grain model was used to simulate the first step in folding, a process which was found to be fast and have a funnel-like landscape (folding from U to N_{iso}). N_{iso}, although native-like, was found to be flexible and consist of an ensemble of closely related structures. Intermediate states observed in the simulations showed a considerable amount of native secondary structure with only the C- or the N-terminus unfolded to varying degrees. Notably, such intermediate states have also been noted in single-molecule pulling experiments[22,23] and other simulations,[26] see section 5.1

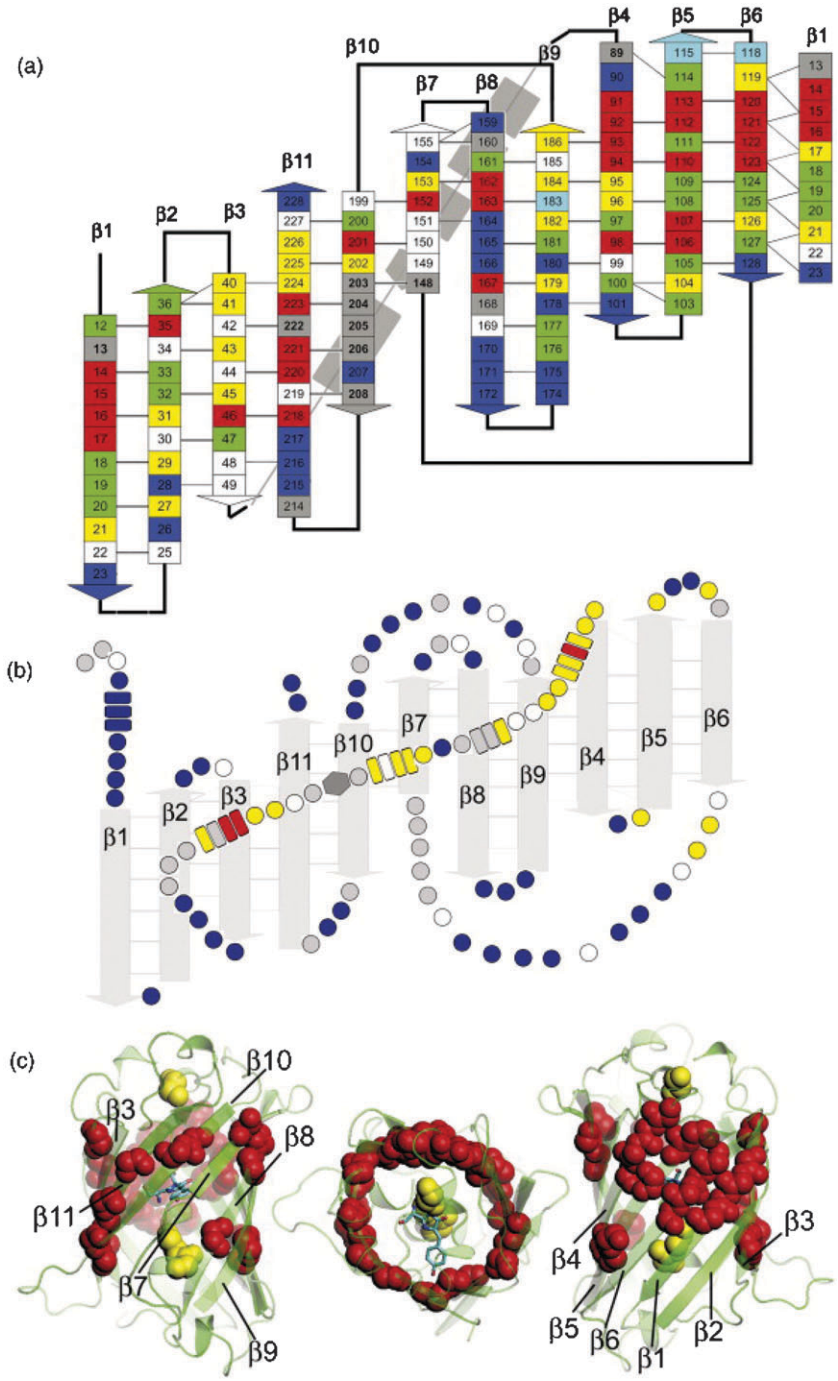


Fig. 3 Results from H/D exchange experiments on trGFPuv. (a) The network of hydrogen bonds in the β-barrel structure. A double line between two residues on two anti-parallel β-strands represents two hydrogen bonds. A single line represents one hydrogen bond. The measured exchange rate constants for the amide groups are classed as follows. Red, very slow (the half-life is longer than one month); yellow, slow; green, intermediate; blue, fast; dark blue, very fast (amide hydrogen exchanged within 20 min); grey, not assigned; white, overlapped peaks in the HSQC spectrum. (b) Amides in regions of random coil (circles) or α-helices (rectangles). Red, very slow; dark blue, very fast; yellow, exchanged within a month; white, overlapped peaks; grey, not assigned. (c) Three-dimensional representation of the position of very slow exchanging residues. Red and yellow balls represent the residues in the β-strands and α-helix which are very slow to exchange, respectively.

for further details. This was supported by results of NMR experiments which showed evidence for structural heterogeneity, particularly near residues adjacent to prolines which have previously been linked to chromophore formation. Both experimental and computational studies pointed to a dual-basin in the energy landscape for folding of sfGFP in which there is the native state (N) and a closely-related N_{iso} state, where the chromophore is not yet in the correct conformation. Following fast initial folding from U to N_{iso} there is a slow rate-limiting search to N, which involves

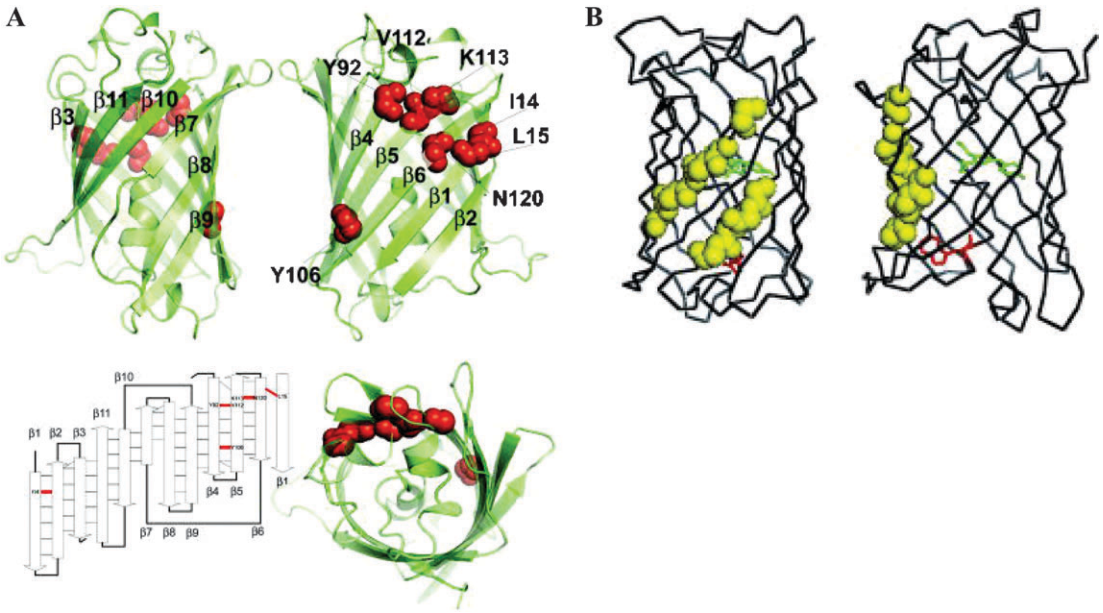


Fig. 4 Results from H/D exchange NMR experiments on trGFPuv. (A) The superstable core of GFP. The three-dimensional structure of GFP showing the positions of all the extremely slow exchanging residues (red balls). Note that they are all located on one side of GFP, and connect β4, β5, β6, β1, and β2. The topology diagram (bottom left) showing the hydrogen bonding by the seven very slow-exchanging amide protons. (B) Three-dimensional representation of the residues with protected amide protons under acid denaturing conditions. Yellow balls: residues with protected hydrogens in the acid-denatured state. Trp57 is shown in red and the chromophore in yellow.

chromophore isomerisation and activation. Recently, further NMR studies on N_{iso} have revealed that several residues in the lid region of sfGFP show structural heterogeneity and that chromophore flexibility leads to mispacking of the protein in this trapped intermediate state.[47]

4. Unfolding and folding kinetics

The folding of GFP was first studied by Reid and Flynn who refolded unfolded, non-fluorescent (chromophore not yet formed) GFP from inclusion bodies and monitored both the folding and chromophore maturation rate.[1] This important paper not only established that the protein had to fold before the cyclisation and oxidation steps required for chromophore formation, but that the folding rate was slow compared to many small proteins, with a half life of some six minutes under the conditions used.[1] The work also proved that chromophore formation was auto-catalytic.

The first, in depth, study on the folding of GFP was published by Kuwajima's group in 2000.[13] The unfolding and refolding of wild-type GFP and GFPuv were compared *in vitro* in order to understand the origins of the improvements in folding and fluorescence observed for GFPuv *in vivo*. The bright fluorescence of the green chromophore was shown to be a sensitive probe of folding and the state of the protein. Unusually slow unfolding and refolding rates were observed for both wild-type and GFPuv, however, little difference in the rates of the two variants was detected. However, wild-type GFP was shown to have a very strong tendency to aggregate whilst GFPuv had a much lower propensity; this was attributed to its lower overall hydrophobicity. In this case, several refolding phases were observed with half lives in the order of 3 and 13 minutes. This work was followed up with another study on the folding of GFPuv which employed multiple probes of structure including chromophore and tryptophan fluorescence and far-UV CD.[18] By using multiple probes, five kinetic refolding phases were observed. These included a very fast (sub-millisecond) step which occurred within the dead-time of the experiments and which was attributed to a non-specific collapse of the polypeptide chain, followed by the population of an on-pathway intermediate state. These experiments also demonstrated that the folding of GFP is dominated and limited by proline isomerisation, perhaps not that surprising given that GFP has ten prolines one of which is in an unfavourable *cis* conformation in the native state. The detailed kinetics experiments were used to establish a complex kinetic scheme for the folding of GFPuv which is shown in Fig. 5A.[18]

At around the same time, two other papers were published where folding rates of different constructs of GFP were reported. In the first case, the β-barrel scaffold of GFP was used to study the effect of cross-strand pairing of different side chains in parallel β-strands.[48] Here, unfolding and folding rates of the different mutants were measured *in vitro*, and compared to the rate of green fluorescence acquisition *in vivo*. Under the conditions used, half lives of approximately three minutes for unfolding and between one and eight minutes for refolding were measured for the different cross-strand pairs. Importantly, the maturation rates of fluorescence observed *in vivo* correlated well with the folding rates measured *in vitro*.[48] In a separate study, Iwai and co-workers used intein technology to cyclise GFP through its N- and C-termini. The unfolding and refolding rates of linear and cyclised GFP were measured using denaturant-jump experiments and the cyclised

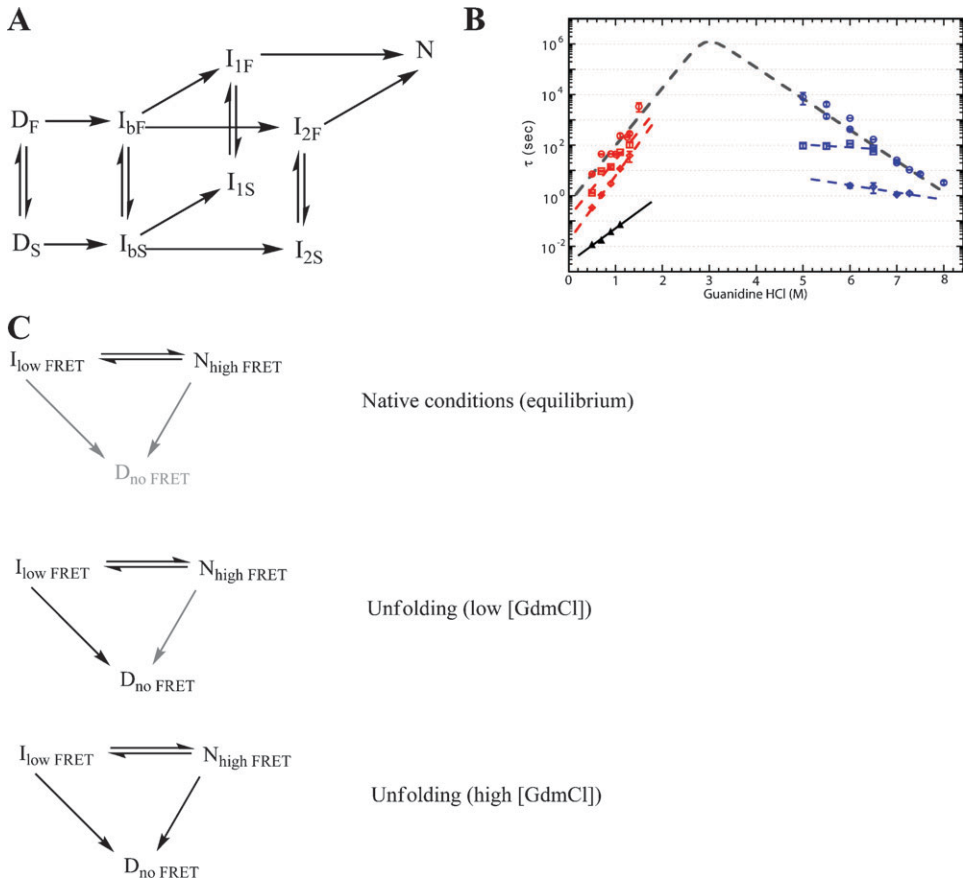


Fig. 5 Unfolding and refolding and kinetics of different forms of GFP. (A) Kinetic scheme for the folding of GFPuv as adapted from ref. 17. (B) Unfolding and refolding kinetics of sfGFP at different concentrations of denaturant taken from ref. 29. (C) Unfolding pathways for an Alexa647-labelled mutant of Citrine under native conditions and at low and high concentrations of GdmCl, taken from ensemble and single-molecule studies, adapted from ref. 32.

protein was found to be more stable than the linear GFP, unfolding at half the rate.[49] At high concentrations of denaturant, half lives for unfolding ranged from 0.1–0.6 minutes, whilst multiple folding phases were observed with half lives in the order of a couple of minutes to 40 minutes reported, consistent with other studies.

Recently, a number of papers have been published that have provided further insight into the un/folding of the complex β-barrel structures of FPs. Studies from our laboratory on the unfolding of trGFPuv confirmed earlier studies that energetic barriers are high and unfolding rates very low in comparison to most small monomeric proteins. The half life for unfolding of GFPuv in water, extrapolated from measurements made in high concentrations of denaturant, is on the order of 500 years.[31] This in some ways explains why FPs are so slow to reach equilibrium in unfolding experiments, in addition to the hysteresis caused by the chromophore (section 3.5). A detailed study of the unfolding and folding kinetics of sfGFP from the Jennings group which appeared at the same time reported similar findings and a half life of unfolding in water of some 28 years under the conditions used.[30] In this case, a comprehensive study of the kinetics as a function of denaturant concentration revealed a lag phase in the folding reaction, followed by three exponential phases the slowest of which was shown to be due to proline isomerisation[30] (Fig. 5B) results very similar to those reported earlier on GFPuv.[18] Interestingly, sfGFP was found to fold an order of magnitude faster than either wild-type GFP or GFPuv.[30]

A recent study by Xie and co-workers on trigger factor (TF) assisted folding of GFPuv has confirmed many of the earlier findings.[29] The unfolding and folding kinetics were found to be complex, intermediate states populated, and rates of folding were accelerated by low concentrations of TF, consistent with proline isomerisation being the rate-limiting step (TF has peptidyl prolyl isomerase activity).

Further insights into the unfolding and folding of variants of GFP have been gained from single-molecule measurements.

5. Single molecule studies of FP unfolding and folding

5.1 Mechanical unfolding

The mechanical strength and stability of monomeric FPs, as well as the mechanical unfolding pathways, have been studied by a number of groups using AFM pulling experiments. The first study published was by the Japanese group of Hara who used AFM to test the mechanical stability of two circular permutants of GFP (cpGFP) where new N- and C-termini

were introduced between residues 144/145 and 172/173.[50] In both cases, the cpGFPs showed reduced mechanical stabilities compared with wild-type GFP and this was attributed to the location of the new termini close to the β-can structure. Fernandez and coworkers have also used circular permutants of enhanced YFP (eYFP) to probe the mechanical unfolding pathway of this protein.[51] In this case, two circular permutants of eYFP were used and all three proteins shown to have similar unfolding peaks corresponding to an initial transition lying close to the termini (most likely displacement of the N-terminal α-helix—see next paragraph for further detail), and one lying approximately halfway through the molecule. It was proposed that this second peak corresponded to the shearing and disruption of β-strands 1 and 6 as these are the only two parallel β-strands in the β-barrel structure and it had been established that the forces required for shearing two neighbouring strands are higher than for unzipping two anti-parallel strands. A point mutation disrupting the interactions between β1 and β6 proved the hypothesis correct.[51] GFP has also been used in other studies to act as a sensitive fluorescent probe in mechanical unfolding experiments. In this case, the mechanical stability of GFP was found to be similar to a fibronectin domain (FN-III).[52]

The most detailed study of the mechanical unfolding of GFP has been undertaken by the Rief group. Their earliest study published in 2004 demonstrated that the mechanical unfolding involves two intermediate states.[23] The first step in unfolding is the unravelling and detachment of the N-terminal seven-residue α-helix away from the β-barrel. This was shown to destabilise the remaining structure of GFP resulting in a thermodynamically unstable state. However, the mechanical unfolding of this state was dependent upon the unfolding activation barrier. A second intermediate state in which a single β-strand had been detached from the β-barrel structure was also observed. Computational approaches have been used to gain further insight into the structures of the intermediate states observed.[53] In particular, these computational methods are required to understand the nature of the second intermediate state as experimental results do not distinguish between mechanisms in which β-strand 1 or 11 is detached. A self-organised polymer model (SOP) was used to simulate the force-induced unfolding of GFP, and the results from these simulations agreed well with the experiments and a model in which the N-terminal α-helix unfolds first followed by the detachment of β-strand 1 in the major unfolding pathway.[53] This work was followed up with a more detailed experimental analysis of GFP variants engineered with disulfide bonds in different positions which were used to effectively lock different regions of structure, in addition to force-unfolding experiments using different points of attachment, thereby altering the direction along which strain propagated during the pulling experiment.[26] Together with the results of coarse-grained SOP simulations, results from these studies established that after the initial detachment of the N-terminal α-helix there is a bifurcation in the unfolding pathway, the major path involving displacement of β-strand 1, then β-strands 2 and 3 leading to a pathway with several intermediate states, and a minor route existing in which β-strand 11 is displaced

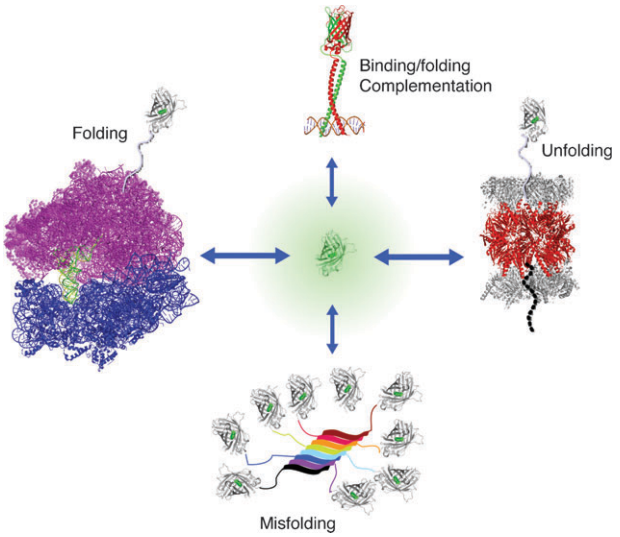


Fig. 6 Overview of the applications of FPs *in vivo* in the study of protein folding, misfolding, unfolding/degradation and visualisation of biomolecular interactions.

following the initial unfolding of the N-terminal helix. In the latter case, the barrel structure was found to flatten and expose 50% of the hydrophobic surface area after displacement of β-strand 11 such that further metastable intermediate states were not observed on this pathway. The fraction of molecules transversing each pathway was found to depend upon mutation or force direction. The rugged energy landscape generated is shown in Fig. 6,[26] and the bifurcated pathway is similar to the results from other single molecule studies of GFP unfolding.[33] However, the intermediate states populated during the force unfolding of GFP appear to be structurally different from those identified by chemical denaturation and acid refolding experiments.[25]

The mechanical unfolding of GFP has also been used by this group to demonstrate how information on protein structure can be obtained using a mechanical triangulation technique.[24] The relative positions of three residues in GFP were predicted from AFM data using cysteine engineered polyprotein forms of the protein.

5.2 Single-molecule fluorescence studies

The chromophore of GFP is sufficiently fluorescent to enable single-molecule fluorescence studies of the protein. A review of single-molecule measurements of GFP is beyond the scope of this review and interested readers are directed towards ref. 54. Here, we will focus on the studies that have used single-molecule measurements to examine the unfolding and folding pathways of GFP and YFP.

An investigation of the switching rates between the anionic (A) and neutral (N) forms of the chromophore in GFPmut2 (S65A; V68L; S72A) established that, whereas normally the switching rates are random, when close to an unfolding transition, they become very regular and only a discrete number of frequencies are observed.[55] The A–N switching rates detected in single molecules were then used to probe the energy landscape around the native state of GFPmut2.[56] Long unfolding time courses were followed by using protein

molecules entrapped in wet nano-porous silica gels. The results from a time course of the A–N switching rate of single molecules in the presence of 5.3 M GdmCl suggested that the unfolding process can only occur along discrete pathways, where the pathway taken is dictated by the initial substate that the molecule was in.[56] Thus, this study provided evidence that the un/folding energy landscape for GFP is rough in comparison with many other proteins, similar to the conclusions of other studies.[30,46,47]

Our group has taken a different approach and used an Alexa647-labelled construct of the yellow fluorescent protein, Citrine, and single-molecule FRET and two-colour coincidence detection (TCCD) to investigate the energy landscape for unfolding of this protein.[33] Even under native conditions, two populations were observed in both FRET and TCCD measurements corresponding to a low- and high-FRET species ($\sim$26% and $\sim$73%, respectively). Control experiments confirmed that these populations were not just due to photophysical effects of the chromophore or dye but could be attributed to different conformational states under native conditions. Thus, it was concluded that the low- and high-FRET states correspond to a locally unfolded intermediate state and the fully folded native state, respectively. A series of single-molecule measurements under equilibrium conditions over a wide range of denaturant concentrations demonstrated that the population of the low-FRET species increased as the population of the high-FRET species decreased with increasing GdmCl concentrations, suggesting a conversion from the high-FRET state into the low-FRET state. Kinetic unfolding experiments were underaken using a nanopipette[57] and manual mixing. At high denaturant concentrations ($>$5.5 M GdmCl), the high-FRET state was observed initially as the predominant species in solution which decreased over time, whilst the low-FRET population increased, remained at a steady level and then ultimately decreased.[33] A careful analysis of the kinetic data revealed that two parallel unfolding pathways were required to explain the data and a triangular reaction scheme was invoked where a fraction of the labelled-YFP molecules unfolded *via* the intermediate state with the low-FRET signal, whilst a fraction unfolded directly to the unfolded state (Fig. 5C).

6. The use of FPs in studying protein folding and misfolding

The characteristic, spontaneous formation of the fluorescent chromophore in GFP and its variants makes them ideal fusion tags for a myriad of *in vivo* biological applications from imaging the cellular localisation of proteins to mapping interactions *in situ* (Fig. 6). However, the tendency of FP variants to aggregate into non-fluorescent species has been a recurrent issue since GFP was first cloned from jellyfish and expressed as a recombinant protein in many different organisms.[58] Protein engineering techniques have therefore been used extensively to improve the folding efficiency and solubility of the protein (see section 2). The aggregation propensity of GFP has been attributed to the complex topology, *i.e.*, high contact order,[59] of the β-barrel structure that in turn is responsible for the high chemical, thermal and mechanical stabilities of FPs (see section 3).

Robust folding of FPs is crucial when they are used as fluorescent fusion tags to study protein stability both *in vitro* and *in vivo*[60–64] (the cellular stability of a protein is likely to be a complex function of its thermodynamic stability in addition to the rate at which it is degraded). In these cases, it is assumed (i) that FPs fold autonomously into natively fluorescent states as fusion tags, and (ii) that fluorescence quenching of FPs reflects unfolding and/or misfolding events that lead to exposure of the chromophore to bulk solvent. In many cases, these assumptions are reasonable and thus the intrinsic fluorescence of FPs is frequently used as a convenient readout for assessing the folding state of FPs as well as the protein targets to which they are tagged.[65] It should be noted, however, that the fluorescence of some FPs is sensitive to the surrounding environment and changing conditions such as pH and halide ion concentrations.[33,66–69] In these cases, changes to the photophysics of the chromophore can occur in the absence of any unfolding event, thereby turning the FP into a "dark state" (see other reviews in this special issue[3,70]). This can cause misleading results and incorrect conclusions. Additionally, the fluorescence properties of the chromophores are dependent on structural factors such as hydrogen bonding and van der Waals contacts, hence the fluorescence of FPs can also be modulated by even small local conformational rearrangements near the chromophore as a result of propagated conformational changes from the tagged proteins. This has been observed in the case of calcium ion binding-induced folding of calmodulins that exhibit marked transitions from a disordered to an ordered state.[71] These particular characteristics have been exploited by Tsien, Miyawaki and others to generate chimeric proteins which can act as pH and ion sensors in living cells.[6,65,68,72,73] Whilst such environment-dependent fluorescence properties have their niche in certain areas, continuous efforts have been made towards generating FP variants with improved folding efficiencies and photophysics under physiological conditions.[15,72,74] These developments are critical when FPs are used as donors and acceptors in FRET measurements *in vivo* where the observed fluorescence is designed to report changes in the distances between donors and acceptors independent of the cellular environment.[65,75] Additionally, it is important for studying protein expression at a single-molecule level where the emissions of individual FP molecules are expected to be quantised with the same emission wavelength.[76]

As outlined above, the fluorescence of FPs is a convenient but not necessarily exclusive parameter for evaluating the folding properties of target proteins, including the FPs themselves. Whilst the fluorescence of FPs is often used as the hallmark of folding, it is essential to employ complementary structural probes when FPs are subject to (un)folding, particularly *in vivo*. Recently, we have used NMR H/D exchange experiments to identify the presence of a superstable core in GFPuv[31,32] and Venus[42] (see section 3.4 for further detail). Our *in vitro* folding data have shown that the N-terminal part of GFPuv remains highly ordered under denaturing conditions and thus we can hypothesise that a folding intermediate with some native-like interactions may be attained during protein translation, *i.e.*, that for GFP, some

co-translational folding can occur.[77–80] While our hypothesis simply assumes that the most stable part of a protein molecule is likely to be the folding nucleus, a recent report by Broude and co-workers has provided evidence that a fluorogenic chromophore can auto-catalytically form within an isolated N-terminal fragment of the enhanced green fluorescent protein (EGFP),[81] lending support to such a hypothesis. Further, a recent single molecule study on a ribosome-bound nascent chain of GFP demonstrated some interesting differences between the *de novo* folding and *in vitro* refolding of GFPuv based on the kinetics of chromophore maturation.[82] This study concluded that the N-terminal region of GFP nascent chains may populate on-pathway intermediate structures that facilitate the subsequent folding process upon release of the chain from the ribosome. Since the catalysis of chromophore formation requires a well-defined conformation maintained by a native-like folding scaffold, this also substantiates our hypothesis.

The maturation of chromophores in FPs involves a rate-limiting oxidation step that typically takes minutes to days depending upon the chromophore, as compared to the folding kinetics which are usually on the timescale of seconds to minutes.[1,13,18,74] The gap between the different reaction timescales, folding *versus* chromophore maturation, highlights some potential issues for the use of FP fluorescence to report on folding of proteins *in vivo*. One commonly used complementary approach for assessing the degree of folding of FPs is a simple solubility assay. By separating the soluble and insoluble fractions of total cell lysates containing over-expressed recombinant proteins, followed by antibody detection, it has been shown that fusion of GFPuv at the N- or C-termini of target proteins can lead to an increased aggregation propensity associated with the co-translational folding process.[83] It has therefore been concluded that multi-domain proteins and those with complex folding topologies, such as FPs, are more likely to aggregate in prokaryotic systems, *e.g.*, *E. coli*, in comparison with eukaryotic systems.[83] This has been attributed to the fact that eukaryotic systems have evolved an expanded repertoire of auxiliary factors, *e.g.*, molecular chaperones, to facilitate *de novo* co-translational folding of proteins. In light of the quest for so-called super-folder FPs that fold well even when fused to poorly folded polypeptide chains,[15] stabilising co-translational folding intermediates may provide a productive alternative strategy and some solutions.

7. The use of complementation and folding of FPs as functional reporters

In a series of recent reports, Waldo and co-workers have used sfGFP to create a split-GFP solubility reporter for the high-throughput evaluation of protein solubility.[84–87] By tagging the last β-strand of sfGFP (β11) to a target protein to create a bait and using a truncated sfGFP that lacks β-strand 11 (sfGFP(1–10)) and which is non-fluorescent, green fluorescence can be restored upon association of the two complementary constructs. If, however, β11 is unavailable for binding and folding with sfGFP(1–10) due to aggregation of the target protein to which it is tagged, there will be no restoration of the green fluorescence. The implications of the experimental design are three-fold: first, sfGFP(1–10) itself is soluble and likely folded; second, the binding of sfGFP(1–10) and β11 is not impeded by the target protein while both are highly soluble; third, the docking of β11 onto sfGFP is sufficient to yield the native structure that gives rise to the fluorescence.

Similar to the split-GFP technology is the so-called bio-molecular fluorescence complementation technique (BiFC)

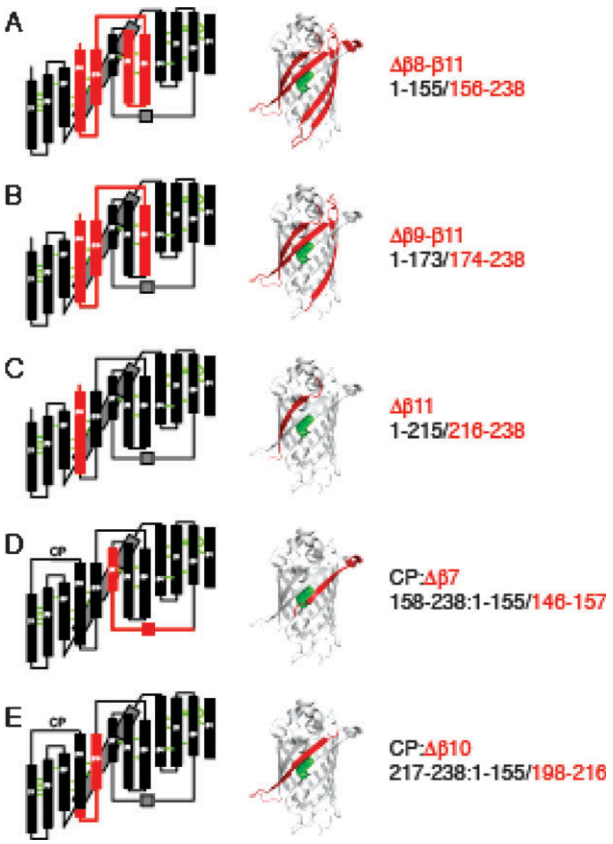


Fig. 7 Constructs of FP variants used in BiFC applications. The left hand panel illustrates schematically the topology of FPs and the eleven β-strands that make up the β-barrel. The stable hydrogen bonds that are formed between strands are indicated by dashed lines. The two N- and C-terminal fragments that constitute the split GFP used in the BiFC applications are coloured in black and red, respectively. The middle panel shows a three-dimensional representation of the GFP constructs where the C-terminal fragment is highlighted in red (the N-terminal fragment is shown in grey and the chromophore as green spheres). The right hand panel provides details of the region comprising each fragment: (A) Construct where the N-terminal region spans residues 1–155 and the C-terminal fragment includes residues 156–238. (B) Construct where the N-terminal region spans residues 1–173 and the C-terminal fragment includes residues 174–238. (C) Construct where the N-terminal region spans residues 1–215 and the C-terminal fragment includes residues 216–238. (D) Circular permutants where the new N- and C-termini have been introduced between the N- and C-termini of β-strand 7 and where β-strand 7 is contained within the complementary fragment. (E) Circular permutants where the new N- and C-termini have been introduced between the N- and C-termini of β-strand 10 and β-strand 10 is contained within the complementary fragment.

that has been pioneered by Kerppola and coworkers and developed to monitor protein–protein interactions *in vivo*. It follows the same principle, that the encounter of two complementary fragments of a FP is sufficient to achieve correct folding to the native state thereby giving rise to fluorescence.[88,89] Using standard molecular biology methods, two split fragments of a FP can be fused to two binding partners and expressed *in vivo*, enabling the direct visualisation of binding events in a variety of cell types and organisms (Fig. 7). The same principle can also be applied *in vitro* with individually purified components. Unlike the split-GFP technology, the two complementary fragments of the FPs used in BiFC are usually similar in size, and separate the protein into two halves at positions 155 or 173, that is at either ends of β7 (Fig. 7). This has been shown to be the least stable β-strand of the barrel structure.[40,90] The original report of BiFC by Hu *et al*. used a heterodimeric leucine zipper DNA binding motif from Jun and Fos as a model system to demonstrate the effectiveness in visualising nuclear localisation of the ternary complex of Jun–Fos–DNA in mammalian COS-1 cell lines.[88] In this case, each of the binding partners, namely Jun and Fos, bears N- or C-terminal fragments of different regions of the FP. Importantly, it is possible to express different FP fragments tagged to different proteins which results in different fluorescence emission depending on the combination of the self-assembled FPs. The wavelengths of the emitted fluorescence are predominantly determined by the sequence composition of and around the chromophore. Such a setup therefore enables differentiation of multiple binding events *in situ*. A large number of applications have since been reported in the literature, including further explorations of combining circular permutation, different truncation sites and different FPs to yield better fragments for BiFC[89,91–96] (Fig. 7). For further details of both the methodology and applications see the review by Kerppola in this special issue.[97]

8. Chaperone-mediated unfolding and degradation

GFP has been used to study both chaperone- and proteosome-mediated degradation pathways. In prokaryotes, there are conserved molecular machines consisting of a chaperone domain, ClpA/ClpC or ClpX, and a protease domain, ClpP—both of which have oligomeric ring-like structures—that unfold, translocate and degrade substrate proteins bearing specific sequences in an ATP-dependent manner.[98,99] Being a mechanically stable β-barrel protein,[22,23,26] GFP has been employed, usually fused to an 11-residue ssrA sequence at the C-terminus, to examine the mechanism by which ClpA/ClpX unfold proteins vectorially from the C-terminus to the N-terminus.[100–103] The order being the reverse of that found in the biogenesis of GFP.[82] Similar to the *de novo* folding studies on GFP variants discussed above, the fluorescence of FPs is used in these studies as a readout of native structure. However, some care must be taken in the interpretation of results from these experiments, as we and many others have shown that the loss of fluorescence does not preclude the presence of persistent structure.

Abbreviations used

GFP	green fluorescent protein
YFP	yellow fluorescent protein
rsGFP	red-shifted green fluorescent protein
FRET	fluorescence resonance energy transfer
TCCD	two-colour coincidence detection
photoCIDNP	photochemically induced dynamic nuclear polarization
BiFC	bimolecular fluorescence complementation
BFP	blue fluorescent protein
CFP	cyan fluorescent protein

References

1 B. G. Reid and G. C. Flynn, Chromophore formation in green fluorescent protein, *Biochemistry*, 1997, **36**, 6786–6791.
2 T. D. Craggs, Fluorescent Proteins: Structure, Folding, and Chromophore Maturation, *Chem. Soc. Rev*., 2009, DOI: 10.1039/b903641p.
3 J. van Thor, Photoreactions and dynamics of the Green Fluorescent Protein, *Chem. Soc. Rev*., 2009, DOI: 10.1039/b820275n.
4 R. Y. Tsien, The green fluorescent protein, *Annu. Rev. Biochem*., 1998, **67**, 509–544.
5 G. S. Waldo, B. M. Standish, J. Berendzen and T. C. Terwilliger, Rapid protein-folding assay using green fluorescent protein, *Nat. Biotechnol*., 1999, **17**, 691–695.
6 G. S. Baird, D. A. Zacharias and R. Y. Tsien, Circular permutation and receptor insertion within green fluorescent proteins, *Proc. Natl. Acad. Sci. U. S. A*., 1999, **96**, 11241–11246.
7 S. Topell and R. Glockshuber, Circular permutation of the green fluorescent protein, *Methods Mol. Biol*., 2002, **183**, 31–48.
8 S. Topell, J. Hennecke and R. Glockshuber, Circularly permuted variants of the green fluorescent protein, *FEBS Lett*., 1999, **457**, 283–289.
9 K. R. Siemering, R. Golbik, R. Sever and J. Haseloff, Mutations that suppress the thermosensitivity of green fluorescent protein, *Curr. Biol*., 1996, **6**, 1653–1663.
10 Y. Kimata, M. Iwaki, C. R. Lim and K. Kohno, A novel mutation which enhances the fluorescence of green fluorescent protein at high temperatures, *Biochem. Biophys. Res. Commun*., 1997, **232**, 69–73.
11 A. Crameri, E. A. Whitehorn, E. Tate and W. P. C. Stemmer, Improved green fluorescent protein by molecular evolution using DNA shuffling, *Nat. Biotechnol*., 1996, **14**, 315–319.
12 R. Battistutta, A. Negro and G. Zanotti, Crystal structure and refolding properties of the mutant F99S/M153T/V163A of the Green Fluorescent Protein, *Proteins: Struct., Funct., Genet*., 2000, **41**, 429–437.
13 H. Fukuda, M. Arai and K. Kuwajima, Folding of green fluorescent protein and the cycle 3 mutant, *Biochemistry*, 2000, **39**, 12025–12032.
14 O. Scholz, A. Thiel, W. Hillen and M. Niederweis, Quantitative analysis of gene expression with an improved green fluorescent protein, *Eur. J. Biochem*., 2000, **267**, 1565–1570.
15 J. D. Pedelacq, S. Cabantous, T. Tran, T. C. Terwilliger and G. S. Waldo, Engineering and characterization of a superfolder green fluorescent protein, *Nat. Biotechnol*., 2006, **24**, 79–88.
16 G. J. Kremers, J. Goedhart, E. B. van Munster and T. W. J. Gadella, Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Forster radius, *Biochemistry*, 2006, **45**, 6570–6580.
17 G. J. Kremers, J. Goedhart, D. J. van den Heuvel, H. C. Gerritsen and T. W. J. Gadella, Improved green and blue fluorescent proteins for expression in bacteria and mammalian cells, *Biochemistry*, 2007, **46**, 3775–3783.
18 S. Enoki, K. Saeki, K. Maki and K. Kuwajima, Acid denaturation and refolding of green fluorescent protein, *Biochemistry*, 2004, **43**, 14238–14248.
19 F. Khan, I. Kuprov, T. D. Craggs, P. J. Hore and S. E. Jackson, F-19 NMR studies of the native and denatured states of green fluorescent protein, *J. Am. Chem. Soc*., 2006, **128**, 10729–10737.

20 P. V. Vrzheshch, N. A. Akovbian, S. D. Varfolomeyev and V. V. Verkhusha, Denaturation and partial renaturation of a tightly tetramerized DsRed protein under mildly acidic conditions, *FEBS Lett*., 2000, **487**, 203–208.
21 H. Herberhold, S. Marchal, R. Lange, C. H. Scheyhing, R. F. Vogel and R. Winter, Characterization of the pressure-induced intermediate and unfolded state of red-shifted green fluorescent protein—A Static and Kinetic FTIR, UV/VIS and Fluorescence Spectroscopy Study, *J. Mol. Biol*., 2003, **330**, 1153–1164.
22 H. Dietz, F. Berkemeier, M. Bertz and M. Rief, Anisotropic deformation response of single protein molecules, *Proc. Natl. Acad. Sci. U. S. A*., 2006, **103**, 12724–12728.
23 H. Dietz and M. Rief, Exploring the energy landscape of GFP by single-molecule mechanical experiments, *Proc. Natl. Acad. Sci. U. S. A*., 2004, **101**, 16192–16197.
24 H. Dietz and M. Rief, Protein structure by mechanical triangulation, *Proc. Natl. Acad. Sci. U. S. A*., 2006, **103**, 1244–1247.
25 M. Bertz, A. Kunfermann and M. Rief, Navigating the Folding Energy Landscape of Green Fluorescent Protein, *Angew. Chem., Int. Ed*., 2008, **47**, 8192–8195.
26 M. Mickler, R. I. Dima, H. Dietz, C. Hyeon, D. Thirumalai and M. Rief, Revealing the bifurcation in the unfolding pathways of GFP by using single-molecule experiments and simulations, *Proc. Natl. Acad. Sci. U. S. A*., 2007, **104**, 20268–20273.
27 O. V. Stepanenko, V. V. Verkhusha, V. I. Kazakov, M. M. Shavlovsky, I. M. Kuznetsova, V. N. Uversky and K. K. Turoverov, Comparative studies on the structure and stability of fluorescent proteins EGFP, zFP506, mRFP1, "dimer2", and DsRed1, *Biochemistry*, 2004, **43**, 14913–14923.
28 B. Wielgus-Kutrowska, M. Narczyk, A. Buszko, A. Bzowska and P. L. Clark, Folding and unfolding of a non-fluorescent mutant of green fluorescent protein, *J. Phys.: Condens. Matter*, 2007, **19**, 285223.
29 J. B. Xie and J. M. Zhou, Trigger factor assisted folding of green fluorescent protein, *Biochemistry*, 2008, **47**, 348–357.
30 B. T. Andrews, A. R. Schoenfish, M. Roy, G. Waldo and P. A. Jennings, The rough energy landscape of superfolder GFP is linked to the chromophore, *J. Mol. Biol*., 2007, **373**, 476–490.
31 J. R. Huang, T. D. Craggs, J. Christodoulou and S. E. Jackson, Stable intermediate states and high energy barriers in the unfolding of GFP, *J. Mol. Biol*., 2007, **370**, 356–371.
32 J. R. Huang, S. T. D. Hsu, J. Christodoulou and S. E. Jackson, The extremely slow-exchanging core and acid-denatured state of green fluorescent protein, *HFSP J*., 2008, **2**, 378–387.
33 A. Orte, T. D. Craggs, S. S. White, S. E. Jackson and D. Klenerman, Evidence of an intermediate and parallel pathways in protein unfolding from single-molecule fluorescence, *J. Am. Chem. Soc*., 2008, **130**, 7898–7907.
34 I. Kuprov, T. D. Craggs, S. E. Jackson and P. J. Hore, Spin relaxation effects in photochemically induced dynamic nuclear polarization spectroscopy of nuclei with strongly anisotropic hyperfine couplings, *J. Am. Chem. Soc*., 2007, **129**, 9004–9013.
35 F. Kahn, K. Stott and S. E. Jackson, ^{1}H, ^{15}N and ^{13}C backbone assignment of the Green Fluorescent Protein (GFP), *J. Biomol. NMR*, 2003, **26**(3), 281–282.
36 J. Georgescu, T. Rehm, J. Wiehler, B. Steipe and T. A. Holak, Backbone H^N, N, C^{α} and C^{β} assignment of the GFPuv mutant, *J. Biomol. NMR*, 2003, **25**(2), 161–162.
37 S.-T. D. Hsu, C. Behrens, L. D. Cabrita and C. M. Dobson, ^{1}H, ^{15}N and ^{13}C assignments of yellow fluorescent protein (YFP) Venus, *Biomol. NMR Assignments*, 2009, **3**(1), 67–72.
38 A. R. Fersht, *Structure and Mechanism in Protein Science: Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman & Co. Ltd, New York, 3rd edn, 2004.
39 S. E. Jackson, How do small single-domain proteins fold?, *Folding Des*., 1998, **3**, R81–R91.
40 M. H. J. Seifert, J. Georgescu, D. Ksiazek, P. Smialowski, T. Rehm, B. Steipe and T. A. Holak, Backbone dynamics of green fluorescent protein and the effect of histidine 148 substitution, *Biochemistry*, 2003, **42**, 2500–2512.
41 S. Enoki, K. Maki, T. Inobe, K. Takahashi, K. Kamagata, T. Oroguchi, H. Nakatani, K. Tomoyori and K. Kuwajima, The equilibrium unfolding intermediate observed at pH 4 and its relationship with the kinetic folding intermediates in green fluorescent protein, *J. Mol. Biol*., 2006, **361**, 969–982.
42 S. T. D. Hsu, G. Blaser, C. Behrens, L. D. Cabrita, C. M. Dobson and S. E. Jackson, Folding studies of venus reveals a strong chloride ion dependence of its yellow fluorescence under mildly acidic conditions, *J. Biol. Chem*., 2009, submitted.
43 C. Scharnagl, M. Reif and J. Friedrich, Stability of proteins: Temperature, pressure and the role of the solvent, *Biochim. Biophys. Acta, Proteins Proteomics*, 2005, **1749**, 187–213.
44 S. W. Englander, Protein folding intermediates and pathways studied by hydrogen exchange, *Annu. Rev. Biophys. Biomol. Struct*., 2000, **29**, 213–238.
45 M. H. Seifert, D. Ksiazek, M. K. Azim, P. Smialowski, N. Budisa and T. A. Holak, Slow exchange in the chromophore of a green fluorescent protein variant, *J. Am. Chem. Soc*., 2002, **124**, 7932–7942.
46 B. T. Andrews, S. Gosavi, J. M. Finke, J. N. Onuchic and P. A. Jennings, The dual-basin landscape in GFP folding, *Proc. Natl. Acad. Sci. U. S. A*., 2008, **105**, 12283–12288.
47 B. T. Andrews, M. Roy and P. A. Jennings, Chromophore packing leads to hysteresis in GFP, *J. Mol. Biol*., 2009, **392**(1), 218–227.
48 J. S. Merkel and L. Regan, Modulating protein folding rates *in vivo* and *in vitro* by side-chain interactions between the parallel beta strands of green fluorescent protein, *J. Biol. Chem*., 2000, **275**, 29200–29206.
49 H. Iwai, A. Lingel and A. Pluckthun, Cyclic green fluorescent protein produced *in vivo* using an artificially split PI-PfuI intein from *Pyrococcus furiosus*, *J. Biol. Chem*., 2001, **276**, 16548–16554.
50 T. Wang, K. Nakajima, T. Kogure, S. Yokokawa, A. Miyawaki and M. Hara, Single-molecule force microscopy of circularly permuted green fluorescent protein, *Jpn. J. Appl. Phys*., 2004, **43**, 5520–5523.
51 R. Perez-Jimenez, S. Garcia-Manyes, S. R. K. Ainavarapu and J. M. Fernandez, Mechanical unfolding pathways of the enhanced yellow fluorescent protein revealed by single molecule force spectroscopy, *J. Biol. Chem*., 2006, **281**, 40010–40014.
52 N. I. Abu-Lail, T. Ohashi, R. L. Clark, H. P. Erickson and S. Zauscher, Understanding the elasticity of fibronectin fibrils: Unfolding strengths of FN-III and GFP domains measured by single molecule force spectroscopy, *Matrix Biol*., 2006, **25**, 175–184.
53 C. Hyeon, R. I. Dima and D. Thirumalai, Pathways and kinetic barriers in mechanical unfolding and refolding of RNA and proteins, *Structure*, 2006, **14**, 1633–1645.
54 D. W. Pierce and R. D. Vale, *Single-molecule fluorescence detection of green fluorescence protein and application to single-protein dynamics*, in *Methods in Cell Biology*, Academic Press Ltd, London, 1999, vol. 58, pp. 49-73.
55 G. Baldini, F. Cannone and G. Chirico, Pre-unfolding resonant oscillations of single green fluorescent protein molecules, *Science*, 2005, **309**, 1096–1100.
56 G. Baldini, F. Cannone, G. Chirico, M. Collini, B. Campanini, S. Bettati and A. Mozzarelli, Evidence of discrete substates and unfolding pathways in green fluorescent protein, *Biophys. J*., 2007, **92**, 1724–1731.
57 S. S. White, S. Balasubramanian, D. Klenerman and L. M. Ying, A simple nanomixer for single-molecule kinetics measurements, *Angew. Chem., Int. Ed*., 2006, **45**, 7540–7543.
58 A. B. Cubitt, R. Heim, S. R. Adams, A. E. Boyd, L. A. Gross and R. Y. Tsien, Understanding, Improving and Using Green Fluorescent Proteins, *Trends Biochem. Sci*., 1995, **20**, 448–455.
59 K. W. Plaxco, K. T. Simons and D. Baker, Contact order, transition state placement and the refolding rates of single domain proteins, *J. Mol. Biol*., 1998, **277**, 985–994.
60 E. Rucker, G. Schneider, K. Steinhauser, R. Lower, J. Hauber and R. H. Stauber, Rapid evaluation and optimization of recombinant protein production using GFP tagging, *Protein Expression Purif*., 2001, **21**, 220–223.
61 N. C. Shaner, G. H. Patterson and M. W. Davidson, Advances in fluorescent protein technology, *J. Cell Sci*., 2007, **120**, 4247–4260.
62 H. Choe, W. H. Li, P. L. Wright, N. Vasilieva, M. Venturi, C. C. Huang, C. Grundner, T. Dorfman, M. B. Zwick, L. P. Wang, E. S. Rosenberg, P. D. Kwong, D. R. Burton, J. E. Robinson, J. G. Sodroski and M. Farzan, Tyrosine sulfation of human antibodies contributes to recognition of the CCR5 binding region of HIV-1 gp120, *Cell*, 2003, **114**, 161–170.

63 H. C. S. Yen, Q. K. Xu, D. M. Chou, Z. M. Zhao and S. J. Elledge, Global Protein Stability Profiling in Mammalian Cells, *Science*, 2008, **322**, 918–923.
64 S. B. VanEngelenburg and A. E. Palmer, Fluorescent biosensors of protein function, *Curr. Opin. Chem. Biol.*, 2008, **12**, 60–65.
65 B. N. G. Giepmans, S. R. Adams, M. H. Ellisman and R. Y. Tsien, Review – The fluorescent toolbox for assessing protein location and function, *Science*, 2006, **312**, 217–224.
66 T. B. McAnaney, W. Zeng, C. F. E. Doe, N. Bhanji, S. Wakelin, D. S. Pearson, P. Abbyad, X. H. Shi, S. G. Boxer and C. R. Bagshaw, Protonation, photobleaching, and photoactivation of yellow fluorescent protein (YFP 10C): A unifying mechanism, *Biochemistry*, 2005, **44**, 5510–5524.
67 D. Arosio, G. Garau, F. Ricci, L. Marchetti, R. Bizzarri, R. Nifosi and F. Beltram, Spectroscopic and structural study of proton and halide ion cooperative binding to GFP, *Biophys. J.*, 2007, **93**, 232–244.
68 S. Jayaraman, P. Haggie, R. M. Wachter, S. J. Remington and A. S. Verkman, Mechanism and cellular applications of a green fluorescent protein-based halide sensor, *J. Biol. Chem.*, 2000, **275**, 6047–6050.
69 M. Kneen, J. Farinas, Y. X. Li and A. S. Verkman, Green fluorescent protein as a noninvasive intracellular pH indicator, *Biophys. J.*, 1998, **74**, 1591–1599.
70 H. E. Seward and C. R. Bagshaw, Photodynamics of Fluorescent Proteins, *Chem. Soc. Rev.*, 2009, DOI: 10.1039/b901355p.
71 C. J. Li, R. Heim, P. Lu, Y. M. Pu, R. Y. Tsien and D. C. Chang, Dynamic redistribution of calmodulin in HeLa cells during cell division as revealed by a GFP-calmodulin fusion protein technique, *J. Cell Sci.*, 1999, **112**, 1567–1577.
72 T. Nagai, S. Yamada, T. Tominaga, M. Ichikawa and A. Miyawaki, Expanded dynamic range of fluorescent indicators for Ca^{2+} by circularly permuted yellow fluorescent proteins, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 10554–10559.
73 F. Metzger, V. Repunte-Canonigo, S. Matsushita, W. Akemann, J. Diez-Garcia, C. S. Ho, T. Iwasato, P. Grandes, S. Itohara, R. H. Joho and T. Knopfel, Transgenic mice expressing a pH and Cl-sensing yellow-fluorescent protein under the control of a potassium channel promoter, *Eur. J. Neurosci.*, 2002, **15**, 40–50.
74 T. Nagai, K. Ibata, E. S. Park, M. Kubota, K. Mikoshiba and A. Miyawaki, A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications, *Nat. Biotechnol.*, 2002, **20**, 87–90.
75 S. Shimozono and A. Miyawaki, *Engineering FRET constructs using CFP and YFP*, in *Fluorescent Proteins*, Elsevier Academic Press Inc, San Diego, 2nd edn, 2008, vol. 85, p. 381.
76 L. Cai, N. Friedman and X. S. Xie, Stochastic protein expression in individual cells at the single molecule level, *Nature*, 2006, **440**, 358–362.
77 A. A. Komar, A pause for thought along the co-translational folding pathway, *Trends Biochem. Sci.*, 2009, **34**, 16–24.
78 B. Hardesty, T. Tsalkova and G. Kramer, Co-translational folding, *Curr. Opin. Struct. Biol.*, 1999, **9**, 111–114.
79 P. L. Clark, Protein folding in the cell: reshaping the folding funnel, *Trends Biochem. Sci.*, 2004, **29**, 527–534.
80 J. R. Huang, Characterising the folding intermediate and denatured state of the large beta-barrel protein GFP, PhD Thesis, University of Cambridge, 2007.
81 V. V. Demidov and N. E. Broude, Profluorescent protein fragments for fast bimolecular fluorescence complementation *in vitro*, *Nat. Protoc.*, 2006, **1**, 714–719.
82 S. Uemura, R. Iizuka, T. Ueno, Y. Shimizu, H. Taguchi, T. Ueda, J. D. Puglisi and T. Funatsu, Single-molecule imaging of full protein synthesis by immobilized ribosomes, *Nucleic Acids Res.*, 2008, **36**, e70.
83 H. C. Chang, C. M. Kaiser, F. U. Hartl and J. M. Barral, De novo folding of GFP fusion proteins: High efficiency in eukaryotes but not in bacteria, *J. Mol. Biol.*, 2005, **353**, 397–409.
84 S. Cabantous, T. C. Terwilliger and G. S. Waldo, Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein, *Nat. Biotechnol.*, 2005, **23**, 102–107.
85 S. Cabantous, J.-D. Pedelacq, B. L. Mark, C. Naranjo, T. C. Terwilliger and G. S. Waldo, Recent advances in GFP folding reporter and split-GFP solubility reporter technologies. Application to improving the folding and solubility of recalcitrant proteins from Mycobacterium tuberculosis, *J. Struct. Funct. Genomics*, 2005, **6**, 113–119.
86 J. D. Pedelacq, E. Piltch, E. C. Liong, J. Berendzen, C. Y. Kim, B. S. Rho, M. S. Park, T. C. Terwilliger and G. S. Waldo, Engineering soluble proteins for structural genomics, *Nat. Biotechnol.*, 2002, **20**, 927–932.
87 P. Listwan, T. C. Terwilliger and G. S. Waldo, Automated, high-throughput platform for protein solubility screening using a split-GFP system, *J. Struct. Funct. Genomics*, 2009, **10**, 47–55.
88 C. D. Hu and T. K. Kerppola, Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis, *Nat. Biotechnol.*, 2003, **21**, 539–545.
89 C.-D. Hu, A. V. Grinberg and T. K. Kerppola, Visualization of protein interactions in living cells using bimolecular fluorescence complementation (BiFC) analysis, *Curr. Protoc. Cell Biol.*, 2006, Chapter 21, Unit 21.3.
90 V. Helms, T. P. Straatsma and J. A. McCammon, Internal dynamics of green fluorescent protein, *J. Phys. Chem. B*, 1999, **103**, 3263–3269.
91 C. Ottmann, M. Weyand, A. Wolf, J. Kuhlmann and C. Ottmann, Applicability of superfolder YFP bimolecular fluorescence complementation *in vitro*, *Biol. Chem.*, 2009, **390**, 81–90.
92 Y. M. Huang and C. Bystroff, Complementation and Reconstitution of Fluorescence from Circularly Permuted and Truncated Green Fluorescent Protein, *Biochemistry*, 2009, **48**, 929–940.
93 M. Sarkar and T. J. Magliery, Engineering a split-GFP reassembly screen to examine RING-domain interactions between BARD1 and BRCA1 mutants observed in cancer patients, *Mol. BioSyst.*, 2008, **4**, 599–605.
94 Y. J. Shyu, H. Liu, X. H. Deng and C. D. Hu, Identification of new fluorescent protein fragments for bimolecular fluorescence complementation analysis under physiological conditions, *BioTechniques*, 2006, **40**, 61–66.
95 Y. J. Shyu, C. D. Suarez and C. D. Hu, Visualization of ternary complexes in living cells by using a BiFC-based FRET assay, *Nat. Protoc.*, 2008, **3**, 1693–1702.
96 T. K. Kerppola, Biomolecular fluorescence complementation (BiFC) analysis as a probe of protein interactions in living cells, *Annu. Rev. Biophys.*, 2008, **37**, 465–487.
97 T. K. Kerppola, Visualization of molecular interactions using bimolecular fluorescence complementation analysis: characteristics of protein fragment complementation, *Chem. Soc. Rev.*, 2009, DOI: 10.1039/b909638h.
98 J. Kirstein, N. Molière, D. A. Dougan and K. Turgay, Structure and function of a novel type of ATP-dependent Clp protease, *Nat. Rev. Microbiol.*, 2009, **7**(8), 589–599.
99 F. I. Andersson, A. Tryggvesson, M. Sharon, A. V. Diemand, M. Classen, C. Best, R. Schmidt, J. Schelin, T. M. Stanne, B. Bukau, C. V. Robinson, S. Witt, A. Mogk and A. K. Clarke, Structure and function of a novel type of ATP-dependent Clp protease, *J. Biol. Chem.*, 2009, **284**(20), 13519–13532.
100 E. U. Weber-Ban, B. G. Reid, A. D. Miranker and A. L. Horwich, Global unfolding of a substrate protein by the Hsp100 chaperone ClpA, *Nature*, 1999, **401**, 90–93.
101 J. Hinnerwisch, W. A. Fenton, K. J. Furtak, G. W. Farr and A. L. Horwich, Loops in the central channel of ClpA chaperone mediate protein binding, unfolding, and translocation, *Cell*, 2005, **121**, 1029–1041.
102 J. R. Hoskins and S. Wickner, Two peptide sequences can function cooperatively to facilitate binding and unfolding by ClpA and degradation by ClpAP, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 909–914.
103 A. Martin, T. A. Baker and R. T. Sauer, Pore loops of the AAA plus ClpX machine grip substrates to drive translocation and unfolding, *Nat. Struct. Mol. Biol.*, 2008, **15**, 1147–1151.